EMBEDDED IMAGE CODING USING WAVELET TRANSFORM

By

D RANGANADHAM

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

JUNE, 1995

EMBEDDED IMAGE CODING USING WAVELET TRANSFORM

A Thesis Submitted in Partial
Fulfilment of the Requirements
for the Degree of
Master of Technology
by

D.Ranganadham

Department of Elecrical Engineering
Indian Institute of Technology Kanpur
India

June, 1995

CERTIFICATE

It is certified that the work contained in the thesis titled 'EMBEDDED IMAGE CODING USING WAVELET TRANSFORM', by *D Ranganadham* has been carried out under my supervision and that this work has not been submitted elsewhere for a degree

6th June, 1995

Dr Sumana Gupta
Associate Professor
Department of Electrical Engineering
IIT, Kanpur

ACKNOWLEDGEMENTS

It is a great pleasure for me to express my deep sense of gratitude to my thesis supervisor, Dr Sumana Gupta for her invaluable guidance and constant help

I have no words to express my thanks to my friend *srinanna* (K Srinivasa Rao) for his invaluable help in coding the algorithm He has been a constant source of inspiration and motivation during all phases of this work I also wish to thank Kalyan for his highly technical advices

Life has been very enjoyable especially in the company of K C S Reddy, T S R , Chaitanya, Ravi Shankar, Ramulu, Sridhar, Jitender Dass, V B R and other telugu friends I am also thankful to my EE batchmates Varada, Venu, Pai and others who helped me to complete my course work

I thank my DOORDARSHAN and SBRA friends Kamat, Prakash, Kali, Manoj, Murty, Burman, Ravi, Paliwal, Rana for making my stay here memorable

Ranganadh

Abstract

Image compression is now essential for applications such as transmission and storage in databases Much of the recent work in image coding has centered on wavelet transforms, which can be used to generate multiresolution of images Image coding techniques using wavelet transform have been shown to achieve high compression ratios while maintaining very good image quality, due to the fact that the edge characteristics of images can be well preserved at low bit rates The aim of the present work is to obtain very high compression ratios at the same time preserving the image quality In order to achieve this an algorithm called EMBEDDED ZEROTREE WAVELET has been implemented The property of this algorithm is that it generates bits in the bit stream in order of importance, so that the decoder can cease decoding at any point in the bit stream The compression algorithm is based on four key concepts

1 Discrete Wavelet Transform which decorrelates the source image very well

2 Zero tree coding which provides significant maps, indicating the position of significant coefficients

3 successive approximation quantization of the significant coefficients

4 Adaptive arithmetic coding which provides a fast and efficient method for entropy coding the strings of symbols and requires no training and prestored tables

The algorithm runs sequentially and stops whenever a desired bit rate is met The result is a hierarchical image compression suitable for embedded coding The reconstructed image quality is dependent on the number of significant coefficients in the encoded bit stream

Contents

List of Figures

Dedicated

to

Doordarshan

Chapter 1

INTRODUCTION

The use of digital images in communications have increased enormously in the last decade It is of considerable contemporary interest to find efficient representations in order to reduce memory required for storage, improve data access rate and reduce bandwidth and/or time required for transmission over the communication channel Hence image compression assumes greater importance All natural images have a large amount of redundant data This redundancy can be classified as

(i) Spatial(due to the correlation between neighboring pixels in an image)

(ii) Spectral (due to the correlation between color planes or spectral bands)

and (iii) Temporal(due to the correlation between neighboring frames in a sequence of images)

In addition to this, there is also some irrelevant data from the observer's point of view This irrelevancy arises because of the limitations and variations of human visual system sensitivity under different stimuli and viewing conditions

The compression schemes available can be classified into two categories They are lossy or lossless schemes respectively In the lossless scheme also referred to as bit preserving compression , the reconstructed image is numerically identical to the original image on the pixel by pixel basis A modest compression ratio of 2 1 can be achieved by this technique[9] The lossy compression on the other hand can achieve higher compression ratio (10 1) with some potentially visible degradations[9]

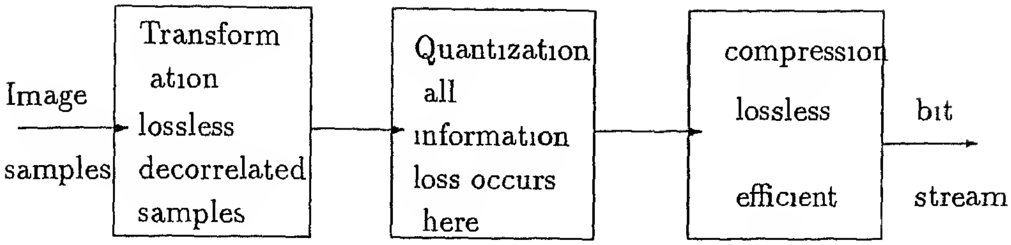


Figure 1 1 A generic transform coder

The general compression scheme comprises of three basic components

(i) Image decomposition or transformations The goal of this process is to decor relate the original image, resulting in the energy being distributed among only a small set of coefficients This is a reversible process

(ii) Image Quantization is a many to one mapping and is a lossy techniques The choice of the type of quantization affects the bit rate and the quality of the recon structed image

(iii) Lossless coding lossless coding of small set of objects resulted form quanti zation to achieve further compression For example zero run length coding , Huffman coding and arithmetic coding

Commonly used techniques in image compression are

(i) $DPCM$(Differential Pulse Code modulation) is a means of predictive coding

(ii) Transform coding Original image is decomposed into a set of basis images Its purpose is to analyze the original image and convert the data into another domain, in which data becomes well behaved and more structured and therefore easier to compress

(iii)Subband coding The image is filtered to create a number of subbands rep resenting various spatial frequency bands of the original full band signal

(iv) Vector Quantization Vector Quantizer could be used directly to encode the image data in a lossy manner

(v) Fractal image compression based on the property of self similarity

The requirement on picture quality and the characteristics of communication ch in

nels and storage media have strong influence on the applied scheme As an exam ple ,TV distribution has preference for high picture quality whereas videophone has preference for worldwide communications with standardized low bit rate channels There are coding schemes for still picture (JPEG)(The joint photographic experts group), H 261 for non interlaced video sequences and MPEG 11 for interlaced video sequences

1 1 JPEG

In the JPEG base true system the input image is divided into disjoint $8X8$ blocks Two dimensional DCT is applied to each block, followed by quantization to reduce the data dynamic range The $2-D$ quantized DCT coefficients are then zigzag scanned into $1-D$ data sequence where the neighboring contiguous zeros grouped together into a run length which can be coded efficiently The Huffman code which assigns short code words to symbols of higher probabilities, is used to code the run length and nonzero coefficients In some of the latter versions of JPEG, arithmetic coding is used in place of Huffman coding which gives better compression

1 2 The objective of thesis

The aim of the current work is

1 To obtain the best image quality for a given bit rate, 2 Accomplishing this task in an embedded fashion, i e in such a way that all encodings of the same image at lower bit rates are embedded in the beginning of the bit stream

The codec shown in fig 1 2 has been used to achieve the above objective

1 3 Organisation of thesis

This thesis has been presented in six chapters In chapter one , we briefly review the conventional methods for image compression In chapter two we briefly review

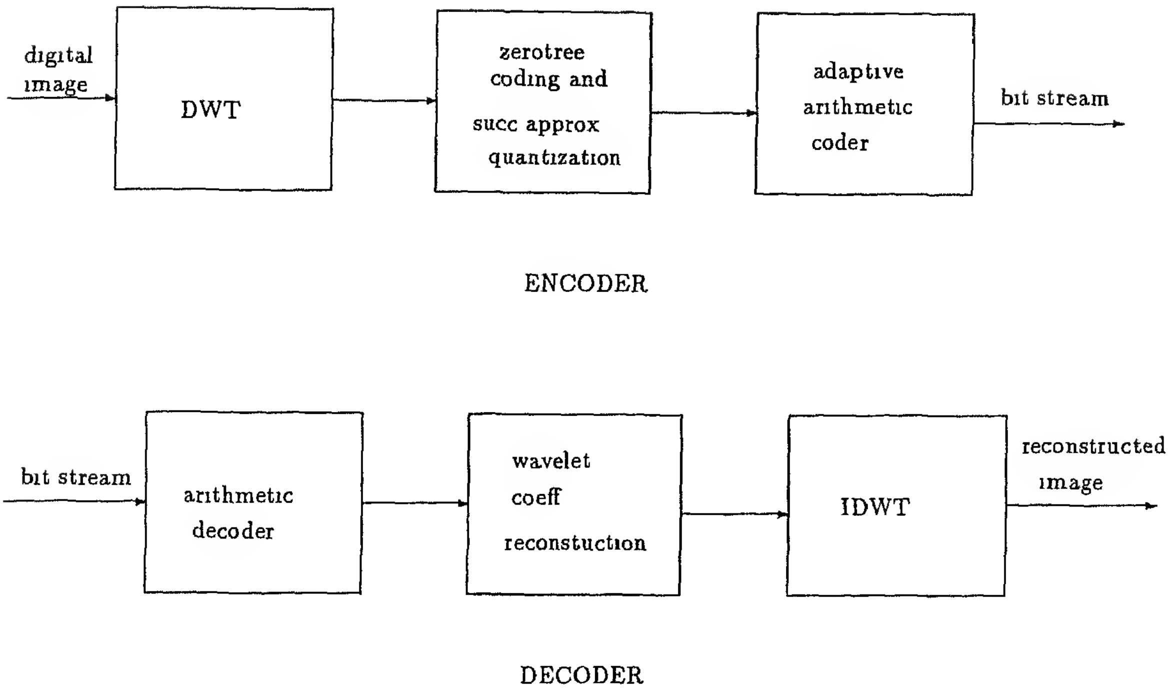


Figure 1 2 Embedded wavelet codec

about the wavelet transform and its extension to $2-D$ Discrete Wavelet transform by which we decompose the digital image into orthogonal basis functions Chapter three deals with embedded zerotree wavelet algorithm by which we remove the insignificant coefficients Chapter four deals with arithmetic coding with fixed and adaptive models to get further compression Chapter five deals with the experimental results and comparison with JPEG and chapter six concludes the thesis and discusses the scope for future work

Chapter 2

DISCRETE WAVELET TRANSFORM

2 0 1 Introduction

An important problem in signal processing is to define a representation that is well adapted for extracting the information content of signals The sharp variations of signal amplitude are generally among the most meaningful features For example the discontinuities of the image intensity provide the contours of the different objects When the signal includes important structures that belong to different scales it is often helpful to reorganize the signal information into a set of detail components of varying size The wavelet transform is a linear operation that decomposes a signal into components that appear at different scales This transformation is based on the convolution of the signal with a dilated filter

2 0 2 Wavelets A brief review

The continuous wavelet transform [9] of signal $s(t)$ is by definition is convolution with a wavelet w(t) dilated by a factor (a) ,

$$W_s(a,b) = a^{-1/2} \int s(t) w((t-b)/a) dt \tag{2 1}$$

$$W(a,b) = a^{-1/2} \int S(w)W(aw)e^{jwb}dw \qquad (2\ 2)$$

Which is equivalent to filtering the signal s(t) with bandpass filter W(aw), whose bandwidth changes according to scale parameter a Clearly, large scale correspond to narrow smoothing filters that represent a global view of the signal s(t) and small scales correspond to wide filters that look into the details of s(t) (i e high frequency components)

Wavelet expansion of the signal s(t) is essentially a decomposition of its frequency content using filters of constant relative bandwidth The signal can be recovered from its wavelet transform coefficients using

$$s(t) = a^{-5/2} \int_a \int_b W_s(a,b)w((t-b)/a)dadb \qquad (2\ 3)$$

assuming that

$$\int_t W(t)dt = 0 \qquad (2\ 4)$$

$$\int_w (W(w)^2/w)dw < \infty \qquad (2\ 5)$$

As it is the case with the Fourier transform, where a signal is expanded in terms of complex exponentials of different frequencies, a wavelet expansion involves dilations of a single wavelet (mother wavelet) The choice of a mother wavelet depends on the application, where a particular wavelet is chosen based on its time and frequency localization

Orthogonality is an important element of wavelet analysis and a mother wavelet is orthogonal to its own dilations and translations Wavelets provide orthonormal basis for expansions of functions that are not of single frequency and are therefore ideal for characterizing signals with discontinuities

2 0 3 Property of wavelets

1 $W(w) = 0$ at w=0 , or equivalently $\int w(t)dt = 0$ i e they have zero dc components

2 They are band pass signals

3 They decay rapidly towards zero with time

Property (1) is a consequence of the admissibility condition of the wavelet the condition that ensures the wavelet transform has an inverse The rapid decay of $w(t)$ is not necessary theoretically for w(t) to be wavelet However, $w(t)$ in practice should have compact support, inorder to have good time localization

2 0 4 Discrete wavelet transform

The wavelet transform parameters can be discretized so that

$$C_{m\,n} = a_0^{-m/2} \int s(t) w(\frac{t - n a_0^m T}{a_0^m}) dt \qquad (2\ 6)$$

$$a = a_0^m \qquad (2\ 7)$$

$$b = n a_0^m T \qquad (2\ 8)$$

and T is sampling period The signal s(t) can be recovered from its expansion coefficients using

$$s(t) = A \sum_m \sum_n C_{m\,n} w \frac{(t - n a_0^m T)}{a_0^m} \qquad (2\ 9)$$

where A is constant

The case where $a_0 = 2$ is known as the dyadic wavelet transform where the signal s(t) is band pass filtered using octave band filters This type of wavelet has the form

$$\psi_{m\,n}(k) = 2^{-m/2} \psi(2^{-m} k - n) \qquad m, n \subset Z \qquad (2\ 10)$$

The discrete wavelet transform (DWT) of discrete time sequence s(k) is essentially a multiresolution characterization of s(k) Generally we take the DWT of a signal that is both time limited and resolution limited A continuous time signal uniformly sampled satisfies this criterion A dyadic discrete wavelet transform is essentially a decomposition of the spectrum of s(k), S(w) into orthogonal subbands defined by

$$\frac{1}{2^j T} <= w <= \frac{1}{2^{j+1} T} \qquad (2\ 11)$$

$j = 1, 2, \quad J$

where T is the sampling period associated with s(k)

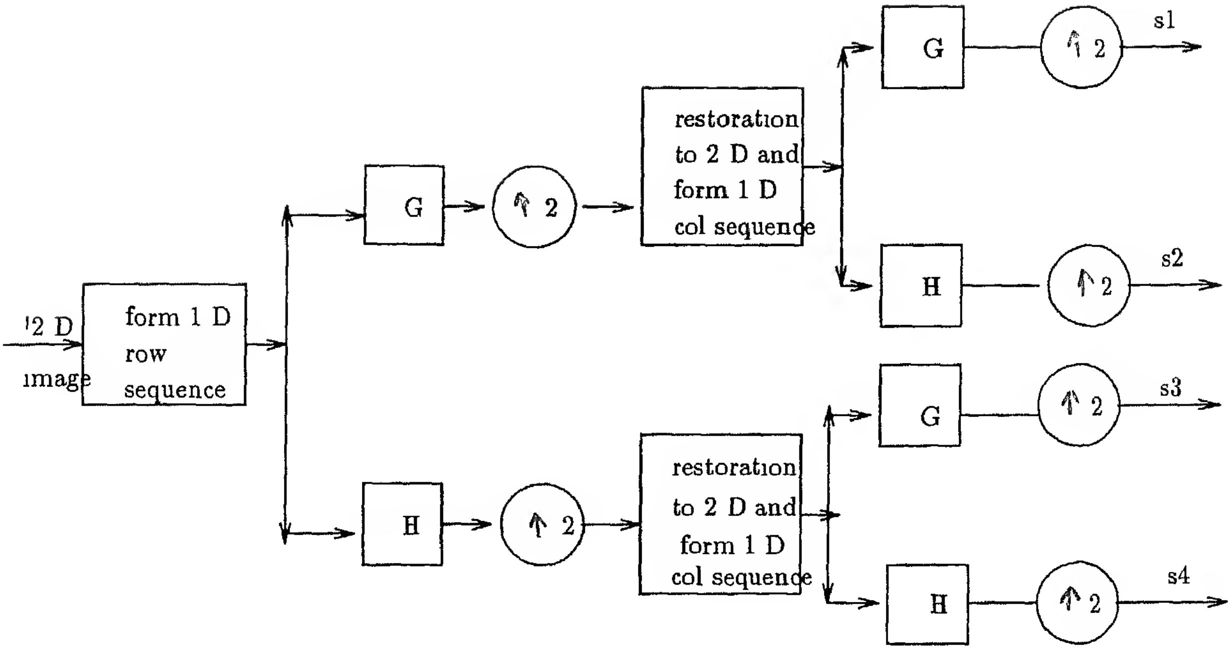


Figure 2 1 2 D MRA decomposition (Multiresolution analysis)

2 0 5 Two Dimensional Wavelets

The idea is to form a 1 D sequence from the 2 D image row sequence, do a 1 D MRA, restore the MRA outputs to a 2 D format and repeat another MRA to the 1 D column sequences The two steps of restoring to a 2 D sequence and forming a 1 D column sequence can be combined efficiently by appropriately selecting the proper points directly from the 1 D MRA outputs As seen from the figure 2 1 after the 1 D row MRA, each low pass and high pass output goes through a 2 D restoration and 1 D column formation process and then move on to another MRA let t_1, t_2 be the 2 D coordinates and L = low pass, H= high pass Then the 2 D separable scaling function is

$$\phi^{(1)}(t_1\ t_2) = \phi(t_1)\phi(t_2)\quad LL \tag{2 12}$$

and the 2 D separable wavelets are

$$\psi^{(2)}(t_1, t_2) = \phi(t_1)\psi(t_2), \qquad LH \tag{2 13}$$

$$\psi^{(3)}(t_1, t_2) = \psi(t_1)\phi(t_2) \qquad HL \tag{2 14}$$

$$\psi^{(4)}(t_1, t_2) = \phi(t_1)\phi(t_2), \qquad HH \tag{ 15}$$

of vanishing moments is defined as

$$\int x^n w(t)dt = 0 \qquad \forall n = 0,1, \quad ,N-1 \tag{2 18}$$

3 Spatial characterization of the scaling function in terms of moments which de termine how $\phi(x)$ evolves with respect to x which allows us to determine the energy concentration of ϕ and provides information on the spatial length or the localization of ϕ This criteria also apply to the wavelet w(t)

4 Characterization of the associated filters To avoid distortion in image pro cessing, the filter $H(\omega)$ associated with the scaling function ϕ must be linear phase or ideally zero phase Indeed, non linear phase filters degrade edges and are more dif ficult to implement than linear phase filters The number of elements making up the impulse response of h(n) must be small in order to limit the number of convolutions operations to be performed in the analysis/reconstruction algorithm It corresponds to wavelets whose support is compact (making the wavelet well localized)

of vanishing moments is defined as

$$\int t^n w(t)dt = 0 \qquad \forall n = 0, 1, \quad , N-1 \tag{2 18}$$

3 Spatial characterization of the scaling function in terms of moments which determine how $\phi(x)$ evolves with respect to x which allows us to determine the energy concentration of ϕ and provides information on the spatial length or the localization of ϕ This criteria also apply to the wavelet w(t)

4 Characterization of the associated filters To avoid distortion in image processing the filter $H(\omega)$ associated with the scaling function ϕ must be linear phase or ideally zero phase Indeed non linear phase filters degrade edges and are more difficult to implement than linear phase filters The number of elements making up the impulse response of h(n) must be small in order to limit the number of convolutions operations to be performed in the analysis/reconstruction algorithm It corresponds to wavelets whose support is compact (making the wavelet well localized)

Chapter 3

ZEROTREE CODING

In image processing, most of the images represent spatial trends, or areas of high statistical spatial correlation However anomalies such as edges or object boundaries, take on a perceptual significance that is far greater than their numerical energy contribution to an image Traditional transform coders, such as those using the DCT, decompose images into a representation in which each coefficient corresponds to a fixed frequency bandwidth, where the bandwidth and spatial area are effectively the same for all coefficients in the representation Edge information tends to disperse so that many non zero coefficients are required to represent edges with good fidelity However, since the edges represent relatively insignificant energy with respect to the entire image, traditional transform coders such as those using DCT, have been fairly successful at medium and high bit rates At extremely low bit rates, however traditional transform coding techniques, such as JPEG tend to allocate too many bits to the trends, and have few bits left over to represent anomalies As a result blocking artifacts often result Wavelet techniques show promise at extremely low bit rates because trends anomalies and information at all scales in between are available Zerotree coding exploits such anomalies across scales

At low bit rates, a large fraction of the bit budget has to be spent to encode the significance map, i e whether a coefficient of 2 D discrete wavelet transform has a zero or non zero quantized value

3 1 Embedded coding

An embedded code[7] represents a sequence of binary decisions that distinguish an image from the null, or all gray, image Since, the embedded code contains all lower rate codes embedded at the beginning of the bit stream, effectively, the bits are placed in order of importance Using an embedded code, an encoder can terminate the encoding at any point thereby allowing a desired bit rate to be met exactly When the desired bit rate is met the encoding simply stops Similarly given a bit stream the decoder can cease decoding at any point and can produce reconstruction corresponding to all lower rate encodings

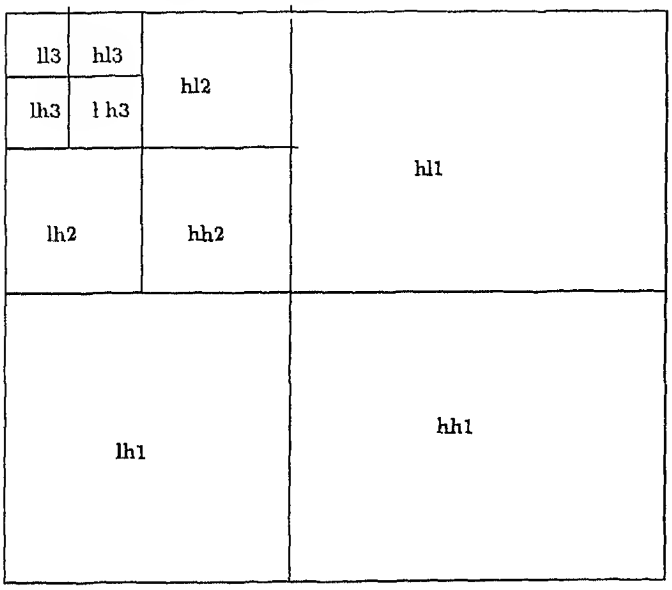


Figure 3 1 3 level wavelet decomposition

3 2 Zerotree coding of wavelet coefficients

In a hierarchical subband system with the exception of the highest frequency sub bands, every coefficient at a given scale can be related to a a set of coefficients at next finer scales The coefficient at the coarse scale will be called the 'parent node and all the coefficients corresponding to the same spatial or temporal location at the next finer scale are called 'child' nodes For a given "parent" node the set of coefficients at all finer scales corresponding to the same location are called 'descendants' Similarly, for a given child, the set of coefficients at all coarser scales of similar orientation

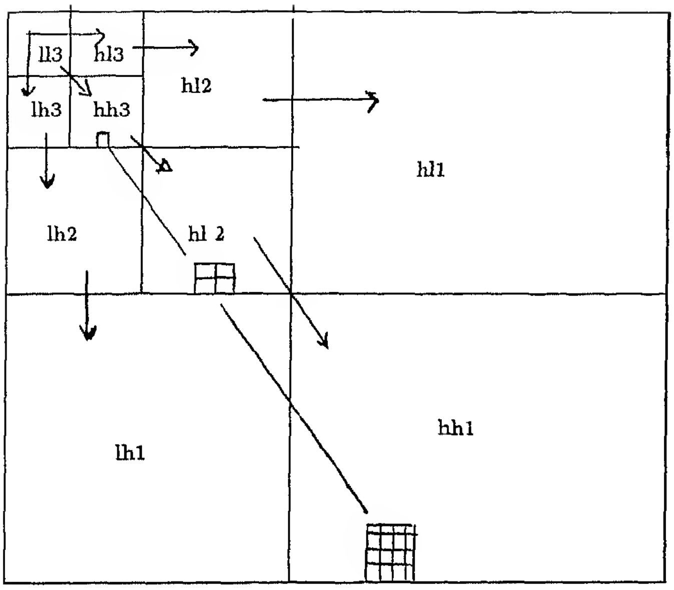


Figure 3 2 Parent child dependencies of subbands Note that the arrow points from the suband of the parents to the subbands of the children The lowest frequency subband is the top left, and the highest frequency subband is at the bottom right Also shown is a wavelet tree consisting of all of the descendents of a single coefficient in suband hh3 The coefficient in hh3 is a zerotree root if it is insignificant and all of its descendents are insignificant

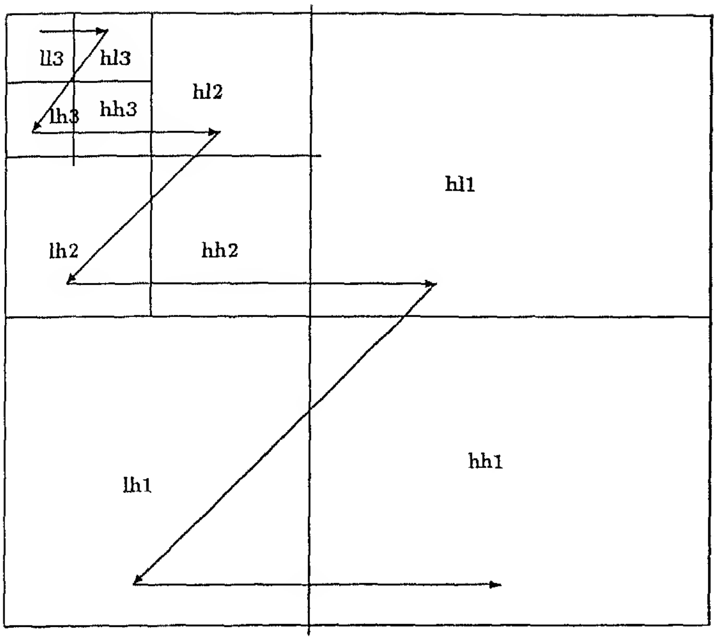


Figure 3 3 Scanning order of the subbands for encoding a significance map Note that all positions in a given subband are scanned before the scan moves to the next subband

corresponding to the same location are called ancestors' For a QMF pyramid sub band decomposition the parent child dependencies are shown in fig 3 2 With the exception of the lowest frequency subband the parent child relationship is defined such that each parent node has three children The scanning of the coefficients is per formed in such a way that no child node is scanned before its parent For a N scale transform the scan begins at lowest frequency subband, denoted as LL_N, and scans subbands HL_N, LH_N, AND HH_N, at which point it moves on the to scale N 1, etc The scanning pattern for a 3 scale QMF PYRAMID is shown in fig 3 3 Note that each coefficient within a given subband is scanned before any coefficient in the next subband

3 2 1 Zerotree data structure

This data structure has four symbols

1 Positive significant (POS) The symbol POS is coded when the coefficient is positive and the magnitude is more than the threshold value,

2 Negative significant (NEG) The symbol NEG is coded when the coefficient is negative and the magnitude is more than the threshold

3 Isolated zero (IZ) The symbol IZ is coded when the magnitude of the coefficient is insignificant, i e less than the threshold, but has at least one significant descendent and

4 Zerotree root (ZTR) The symbol ZTR is coded when the coefficient and all its descendants are insignificant

The zerotree is based on the hypothesis that a wavelet coefficient at a coarse scale is insignificant with respect to a given threshold T, then all wavelet coefficients in the same spatial location at finer scales are likely to be insignificant with respect to T Empirical evidence suggests that this hypothesis is true When encoding the finest scale coefficients, since coefficients have no children, the symbols in the string come from a 3 symbol alphabet, whereby the zerotree symbol is not used The flow chart for the decisions made at each coefficient are shown in fig 3 4

Zerotree coding reduces the cost of encoding the significance map using self similarity Even though the image has been transformed using a decorrelating transform the occurs of insignificant events are not independent events More traditional techniques employing transform coding typically encode the binary map via some form of run length coding Unlike the zerotree symbol, which is single terminating symbol and applies to all tree depths, run length encoding requires a symbol for each run length which must be encoded A technique that is closer in spirit to the zerotrees is the end of block(EOB) used in JPEG, which is also a terminating symbol indicating that all remaining DCT coefficients in the block are quantized to zero To see why zerotree may provide an advantage over EOB symbols consider that a zerotree represents the insignificance information in a given orientation over an approximately square spatial area at all finer scales upto and including the scale of the zerotree root Because the wavelet transform is a hierarchical representation varying the scale in which a zerotree root occurs automatically adapts the spatial area over which the insignificance is represented The EOB symbol however always represent insignificance over same spatial area although the number of frequency bands within this spatial area varies Given a fixed block size such as 8X8, there is exactly one scale in the wavelet transform in which if a zerotree root is found at that scale, it corresponds to same spatial area as the block of the DCT If a zerotree root can be

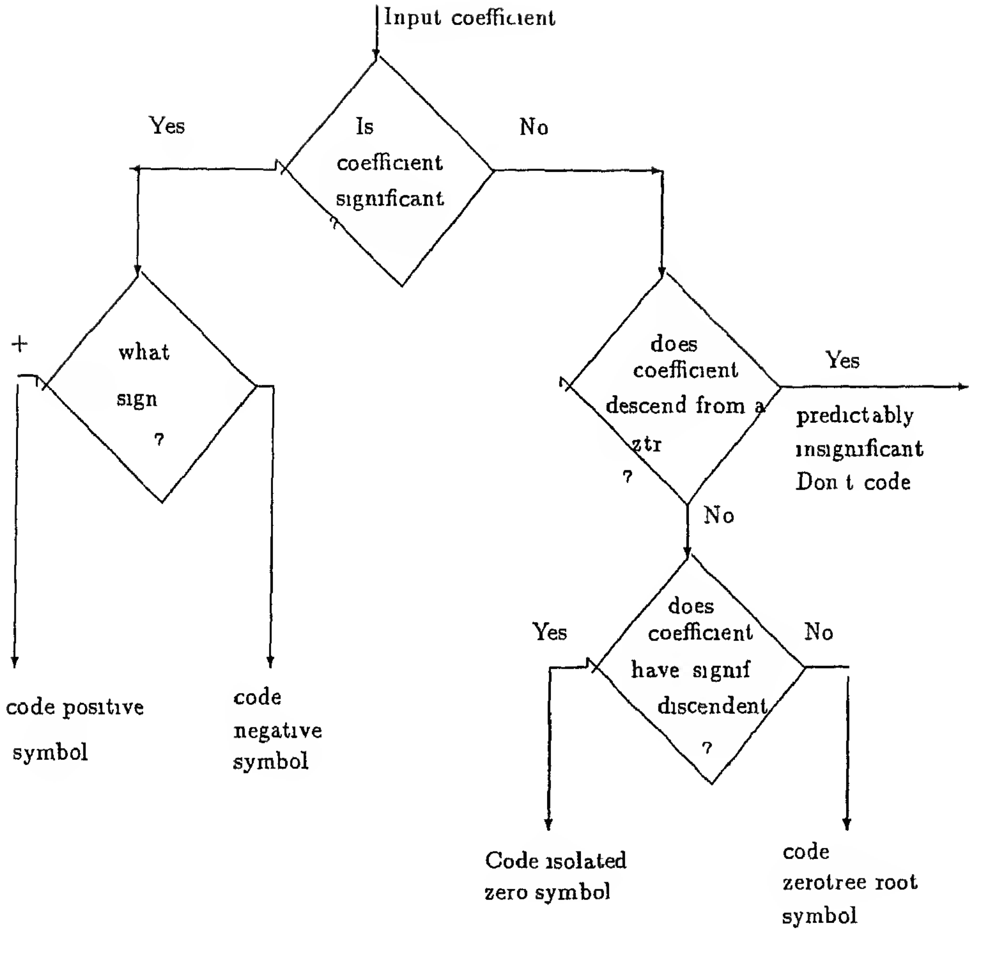


Figure 3 4 Flow chart for encoding a coefficient of the significance map

found at the coarser scale, then the insignificance pertaining to that orientation can be predicted over a large area Zerotree approach can isolate interesting non zero details by immediately eliminating large insignificant region from consideration In this zerotree approach [7] the focus is on reducing the cost of encoding the significant map so that for a given bit budget more bits are available to encode expensive significant coefficients In practice, a large fraction of the insignificant coefficients are efficiently encoded as part of the zerotree

3 3 Successive approximation quantization

In the previous section we described a method of encoding significance maps of wavelet coefficients that, at least empirically seems to consistently produce a code with a lower bit rate than either the empirical first order entropy, or a run length code of the significance map

To perform the embedded coding, successive approximation quantization is applied The successive approximation quantization sequentially applies a sequence of thresholds $T_0, T_1, \quad , T_{N-1}$ to determine significance, where the thresholds are chosen so that $T = T_{\frac{-1}{2}}$

A wavelet coefficient x is said to be insignificant with respect to a given threshold T if $|x| < T$ The initial threshold T_0 is chosen so that $|X_j| < 2T_o$ for all transform coefficients x_j

During the encoding (and decoding) two link lists, one for dominant pass and the other for subordinate pass, of wavelet coefficients are maintained At any point in the process the dominant list keeps track of the coefficients that have not been found to be significant in the same relative order as the initial scan This scan is such that the subbands are ordered, and within each subband the set of coefficients are ordered Thus using the ordering of subands shown in fig 3 3 The subordinate list contains the magnitudes of those coefficients that have been found to be significant For each threshold each list is scanned once

During a dominant pass coefficients that have not been found to be significant in the previous scan are compared to the threshold T to determine the significance and if significant, their sign This significance map is then zerotree coded Each time a

coefficient is encoded as significant, (positive or negative), its magnitude is appended to the subordinate list The coefficient that had been determined to be significant in the previous scan is considered as insignificant in all the following dominant passes so that it does not prevent the occurrence of a zerotree in future dominant passes at smaller thresholds

During a subordinate pass, the width of the effective quantizer step size which defines an uncertainty interval for the true magnitude of the coefficient is halved For each magnitude on the subordinate list, this refinement can be encoded using binary alphabet with a 1 symbol indicating that the true value falls in the upper half of the old uncertainty interval and a 0 symbol indicating the lower half The string of symbols from this binary alphabet that is generated during a subordinate pass is then entropy coded

This process continues alternately between dominant passes and subordinate passes where the threshold is halved before each dominant pass

In the decoding operation, each decoded symbol, both during a dominant and a subordinate pass refines and reduced the width of the uncertainty interval in which the true value of the coefficient (or coefficients in the case of a zerotree root) may occur The reconstruction value can be any where in that uncertainty interval For minimum mean square error distortion, one could use the centroid of the uncertainty region using some model for the PDF of the coefficients However, a practical ap proach, that is used in the experiments, which is also MINMAX optimal, is to simply use the center of the uncertainty interval as the reconstruction value

The encoding stops when some desired terminating condition is met such as when the bit budget is exhausted

3 3 1 Order of importance of bits

Although importance is a subjective term, the order of processing used in EZW algorithm implicitly defines ordering of importance

The primary determination of ordering importance is the numerical precision of the coefficients This is due to the fact that the uncertainty intervals for the magnitude of all the coefficients are refined to the same precision before the uncertainty interval

for any coefficient is refined further

The second factor in the determination of importance is magnitude Importance by magnitude manifests itself during a dominant pass because prior to the pass all coefficients are insignificant and presumes to be zero When they are found to be significant they are all assumed to have the same magnitude which are greater than the magnitude of those coefficients that remain insignificant

The third factor, scale, manifests itself in the a priori ordering of the subbands of the initial dominant list Until the significance of the magnitude of the coefficient is discovered during a dominant pass, coefficients in the coarser scales are tested for significance before coefficients in the finer scales This is consistent with the prioritization of the decoder's version of magnitude since for all coefficients not yet found to be significant the magnitude is presumed to be zero

The final factor, spatial location, merely implies that the coefficients that can not yet be distinguished by the decoder in terms of either precision, magnitude, or scale, have their relative importance determined arbitrarily by the initial scanning order of the subband containing the coefficients

Since a discrete wavelet transform is an invertible representation of an image, a distortion function defined in the wavelet transform domain is also a distortion function defined on the image Since minimizing the widths of uncertainty intervals minimizes the largest possible errors, artifacts, which result from numerical errors large enough to exceed perceptible thresholds are minimized

3 4 Simple example

In this section, a simple example will be used to highlight the order of operations used in the EMBEDDED ZEROTREE algorithm Consider a simple 3 scale wavelet transform of an 8x8 image The array of values is shown in fig 3 5 Since the largest coefficient magnitude 63 which is greater than the threshold 32 and is positive so a positive symbol is generated After decoding this symbol, the decoder knows the coefficient in the interval [32, 64) whose center is 48

2 Even though the coefficient 31 is insignificant with respect to threshold 32, it

63	34	49	10	7	13	12	7
31	23	14	13	3	4	6	1
15	14	3	12	5	7	3	9
9	7	14	8	4	2	3	2
5	9	1	47	4	6	2	2
3	0	3	2	3	2	0	4
2	3	6	4	3	6	3	6
5	11	5	6	0	3	4	4

Figure 3 5 Example of a 3 scale wavelet transform of an 8X8 image

has significant descendant two generations down in subband LH1 with magnitude 47 Thus the symbol for isolated zero is generated

TABLE 1

Processing of first dominant pass at threshold T=32
Symbols are POS for positive significant, NEG for negative significant,
IZ for isolated zero, ZTR for zerotree root, and Z for zero
The reconstruction magnitudes are taken as the center of the uncertainty interval

comment	subband value	coefficient value	symbol	reconstructed
(1)	LL3	63	POS	48
	HL3	34	NEG	48
(2)	LH3	31	IZ	0
(3)	HH3	23	ZTR	0
	HL2	49	POS	48
(4)	HL2	10	ZTR	0
	HL2	14	ZTR	0
	HL2	13	ZTR	0
	LH2	15	ZTR	0
(5)	LH2	14	IZ	0
	LH2	9	ZTR	0
	LH2	7	ZTR	0
(6)	HL1	7	Z	0
(4)	HL1	13	Z	0
	HL1	3	Z	0
	HL1	4	Z	0
	LH1	1	Z	0
(7)	LH1	47	POS	48
	LH1	3	Z	0
	LH1	2	Z	0

3 The magnitude 23 is less than 32 and all descendants which include (3, 12, 14, 8) in subband HH2 and all coefficients in subband HH1 are insignificant A zero tree

symbol is generated, and no symbol will be generated for any coefficients in subbands HH2 and HH1 during the current dominant pass

4 The magnitude 10 is less than 32 and all descendants (12 7, 6 , 1) also have magnitude less than 32 Thus a zerotree symbol is generated

5 The magnitude 14 is insignificant with respect to 32 Its children are (1, 47, 3, 2) Since its child with magnitude 47 is significant, an isolated zero symbol is generated

6 Note that no symbols were generated from subband HH2 which would ordinarily precede subband HL1 in the scan Also note that since subband HL1 has no descendents the entropy coding can resume using a 3 symbol alphabet where the IZ and ZTR symbols are merged into Z (zero) symbol

TABLE 2

Processing of the first subordinate pass
Magnitudes are partitioned into the uncertainty intervals[32,48)
and [48 64) with symbols 0 and 1 respectively

coefficient magnitude	symbol magnitude	reconstructed
63	1	56
34	0	40
49	1	56
47	0	40

7 The magnitude 47 is significant with respect to 32 Note that for the future dominant passes, this position will be replaced with the value 0 so that the next dominant pass at threshold 16 the parent of this coefficient, which has mean square can be coded using zerotree root symbol

During the first dominant pass, which used a threshold of 32, four significant coefficients were identified These coefficients will be refined during the first subordinate

pass Prior to the first subordinate pass, the uncertainty interval for the magnitude of all the significant coefficients in the interval[32, 64) The first subordinate pass will refine these magnitudes and identify them as being either in the interval [32, 48) which will be encoded with the symbol 0, or in the interval [48, 64) which will be encoded with the symbol 1 Thus the decision boundary is magnitude 48 It is no coincidence that these symbols are exactly the first bit to the right of the MSBD in the binary representation of the magnitudes The order of operations in the first subordinate pass is illustrated in the fig

The first entry has magnitude 63 and is replaced in the upper interval whose center is 56 The next entry has magnitude 34 which places it in the lower interval The third entry 49 is in the upper interval, and the fourth entry 47 is in the lower interval Note that in the case of 47, using the center of the uncertainty interval as the reconstruction value when the reconstruction value is changed from 48 to 40 the reconstruction error actually increases from 1 to 7 Nevertheless, the uncertainty interval for this coefficient decreases from width 32 to width 16 At the conclusion of the processing of the entries on the subordinate list corresponding to the uncertainty interval [32, 64), these magnitudes are reordered for future subordinate passes in the order (63, 49 34,47) Note that 49 is moved ahead of 34 because from the decoder point of view , the reconstruction values 56 and 40 are distinguishable However the magnitude 34 remains ahead of magnitude 47 because as far as the decoder can tell both have magnitude 40, and the initial order, which is based first on importance by scale has 34 prior to 47

The process continues on to the second dominant pass the new threshold of 16 During this pass, only those coefficients not yet found to be significant are scanned Additionally those coefficients previously found to be significant are treated as zero for the purpose of determining if a zerotree exists Thus the second dominant pass consists of encoding the coefficient 31 in subband LH3 negative significant the coefficient 23 in subband HH3 positive significant the three coefficients in subband HL2 that have not been previously found to be significant (10 14 13) are each encoded as zerotree roots as are all four coefficients in subband LH2 and all four coefficients in subband HH2 The second dominant pass terminates at this point since all other coefficients are predictably insignificant The subordinate list now contains, in order the magnitudes (63, 49, 34, 47, 31,23) which prior to this subordinate pass represent the three uncertainty interval,s [48, 64), [32, 48) and [16, 31) each having equal width

16 The processing will refine each magnitude by creating two uncertainty intervals for each of the three current uncertainty intervals At the end of the second subordinate pass the order of the magnitude is (63, 49, 47, 34, 31, 23), since at this point, the decoder could have identified 34 and 47 as being in different intervals Using the center of the uncertainty interval as the reconstruction value the decoder lists the magnitudes as (60, 52, 44, 36, 28, 20) The processing continues alternating between dominant and subordinate passes and can stop at any time

Chapter 4

ARITHMETIC CODING

Arithmetic coding is a lossless compression technique that produces an encoded string for an input string of symbols and a model This encoded string represents a fractional value R for the range $0 <= R < 1$

Arithmetic coding [3] is superior to the well known Huffman method in many respects It represents information as compactly as Huffman code It is known that if each symbol in the input string is represented as an integral number of bits in the encoding, then Huffman coding achieves "minimum redundancy' In other words it performs optimally if all symbol probabilities are integral powers of 1/2 But in practice this is normally not so Arithmetic coding dispenses with this restriction that each symbol be represented as an integral number of bits

In arithmetic coding , a message is represented by an interval of real numbers between 0 and 1 As the message becomes longer, the interval needed to represent it becomes smaller and the number of bits needed to specify that interval grows Successive symbols of the message reduce the size of interval in accordance with the symbol probabilities generated by the model The more likely symbols reduce the range by smaller amounts as compared to unlikely symbols and add fewer bits to the message Before transmitting a message the range of the message in the entire interval is [0 1), and the half open interval is denoted by $0 <= x < 1$ As each symbol is processed, the range is narrowed to the portion allocated to the symbol

For example, suppose the alphabet is $a,e,i,ou,'$ and a fixed model is used with

probabilities shown in Table 1

TABLE 1

symbol	probability	range
a	2	[0 0 2)
e	3	[0 2,0 5)
i	1	[0 5 0 6)
o	2	[0 6, 0 8)
u	1	[0 8, 0 9)
!	1	[0 9, 1 0)

Imagine transmitting the message *eaii!* Initially both encoder and decoder know that the range is [0, 1) After seeing the first symbol, *e*, the encoder narrows it to [0 2, 0 5), the range the model allocates to this symbol The second symbol, *a*, will narrow this new range to the first one fifth of it, since a has been allocated [0 0 2) This produce [0 2, 0 26) Since the previous range was 0 3 units long and one fifth of that is 0 06 The next symbol, *i*, is allowed [0 5 0 6) which when applied to [0 2, 0 26) gives the smaller range [0 23, 0 236) proceeding in this way, the encoding message builds up as follows

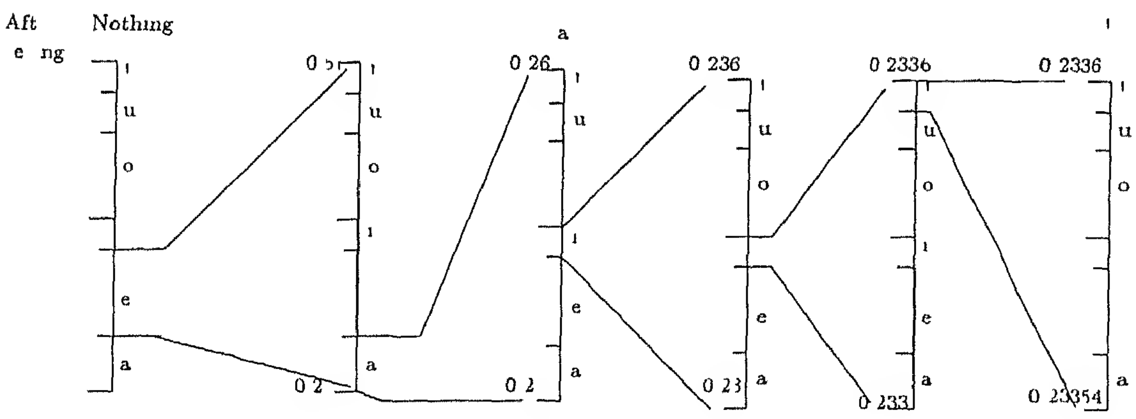


Figure 1 1 Representation of the Arithmetic coding process with the interval scaled up at each stage

The main compression system consists of the model and the encoder In arithmetic

coding the encoder is separate from the model

4 0 1 Models for arithmetic coding

We can divide models for the arithmetic coding into two categories 1 Fixed model and 2 Adaptive model

In the fixed model, frequency of symbol occurrences are taken from sample text

In the adaptive model, frequencies are initialized to some value during encoding, and these frequencies are updated on the basis of symbol frequencies observed in the input string

Arithmetic encoder operates successively on each data symbol determines the context(i e, which relative frequency distribution applies to the current event) and generates the code string Usually, first order Markov model is used to determine the context It takes previous symbol as the context for the current symbol The property used for the coding method is first in first out (FIFO), as it allows for adapting to the statistics of the data string

To represent the magnitude R of the encoded string in the interval [0, 1) great precision is required Fortunately this magnitude need not be given all at once At any stage the upper and lower bounds for R are available as a finite no of digits These digits are left shifted as they become identical and new digits are brought at the low significant end

4 0 2 Arithmetic coding in the context of Zerotree coding

Note that the particular alphabet used by the arithmetic coder at any given time contains either 2 3 or 4 symbols depending on whether the encoding is for subordinate pass or a dominant pass with no zerotree root symbol, or a dominant pass with a zerotree root symbol There is advantage in adapting the arithmetic coder Since there are never more than four symbols all of the possibilities typically occur with in a reasonably measurable frequency This allows an adaptation algorithm with a short memory to learn quickly and track continuouslychanges in symbol probabilities This adaptivity accounts for some of the effectiveness of the overall algorithm On the other

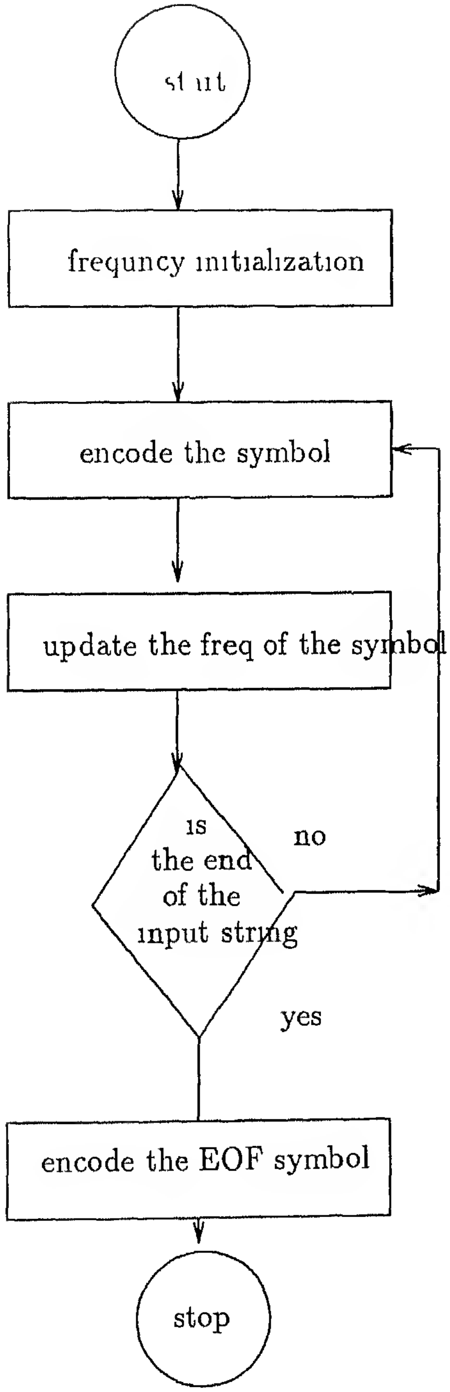


Figure 4 2 Encoding

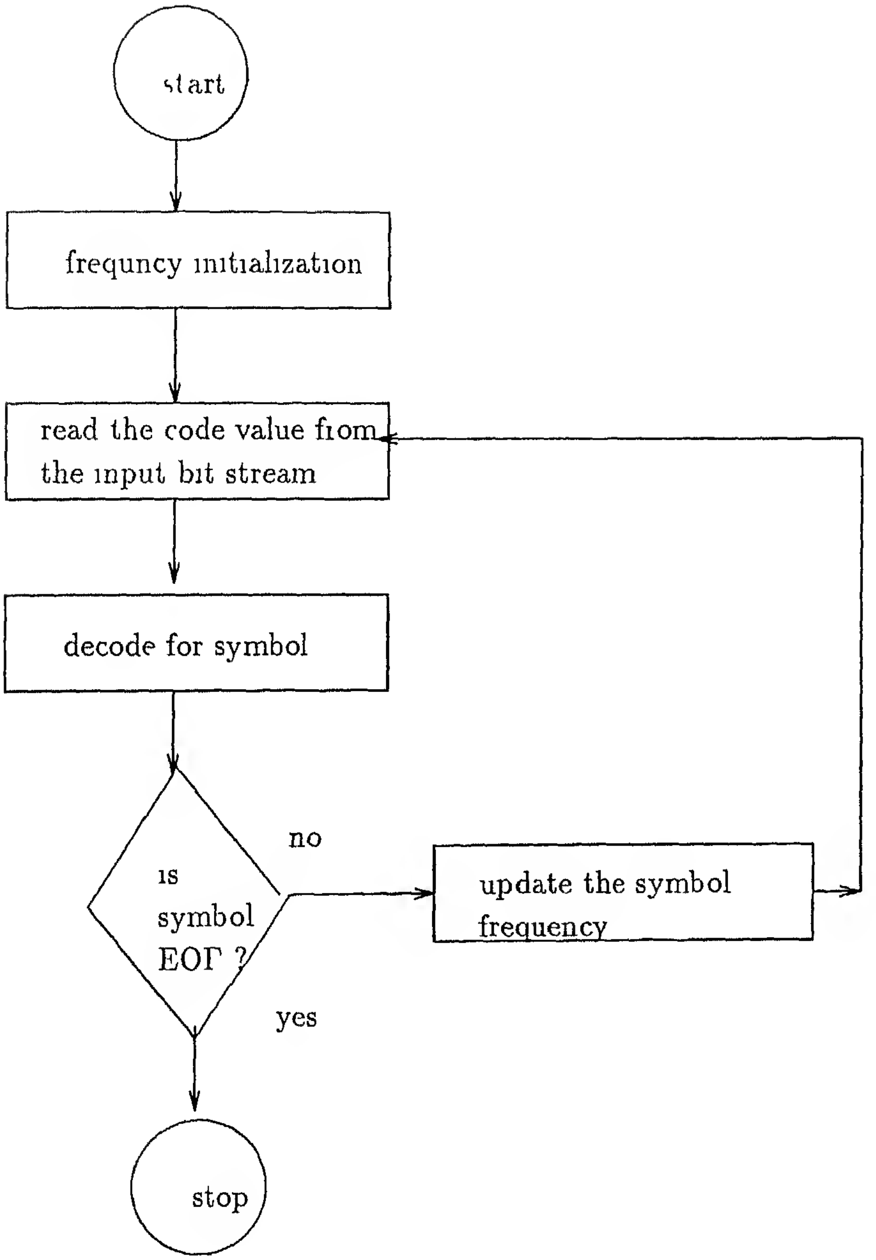


Figure 4 3 Decoding

hand in case of algorithms that do not use successive approximation, several events are needed before an adaptive entropy coder can reliably estimate the probabilities of unlikely symbols

Once the type of model (adaptive model for the present case) is fixed for arithmetic coding, maximum frequency count is the critical parameter for the coding, because it affects underflow, overflow and learning rate for adaptation Arithmetic coding works by scaling the cumulative probabilities given by the model onto the interval [low high] for each character encoded If they are very close together, then there is possibility of mapping different symbols in the same interval Therefore the interval should at least be as large as possible It should not be too large to cause overflow Learning rate for adaptation is inversely proportional to the maximum frequency count Again very small symbol set from zerotree coding is advantageous in choosing maximum histogram count A maximum frequency count of 256 is used taking into account all the factors given above

Chapter 5

EXPERIMENTAL RESULTS

The method implemented here is by nature a lossy method, as some of the wavelet coefficients are eliminated However, by allowing more passes in the successive approximation quantization, i e for higher bit rates, the distortion caused by the removal of information is minimized All experiments were performed by encoding and decoding the actual bit stream to verify the correctness of the algorithm After a 8 byte header the entire bit stream is arithmetically encoded by a single arithmetic coder with an adaptive model[3] The 8 byte header contains 1 No of wavelet scales 2 The dimension of the image 3 The initial threshold and 4 Mantissa

Note that after the header file, there is no overhead except for an extra symbol for end of bit stream, which is always maintained at minimum probability This extra symbol is not needed for storage on computer medium if end of a file can be detected

The simulations are performed on 512 by 512 black and white Lenna and Baboon images respectively The intensity of each pixel is coded on 256 grey levels (8 bpp) The coding results are summarized for Lenna and Baboon in table 9 and table 10 respectively

The filters used to compute the discrete wavelet transform in this thesis is based on symmetric quadrature mirror filters of length (4, 8, 10, 12 20) given by Daubechies[8] We have got the best reconstructed image quality using filter length of 4 Six scales of QMF pyramid were used We can go for higher level wavelet decomposition but there is not much improvement in the reconstructed picture quality and it takes more

time if we go for higher level decomposition Similar results are shown for 256 by 256 Lenna Baboon, and Nutan images in table 6, table 7 and table 8 respectively For 256 by 256 images five levels of wavelet decomposition is used

The fidelity criteria used in most of the DCT based image coding is MSE (mean square error) and PSNR (peak signal to noise ratio) The numerical evaluation of the coder's performance is achieved by computing the PSNR between the original image and the coded image Generally small values of MSE correspond to perceptually high quality reconstructed image

$$MSE = (N_1 N_2)^{-1} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} (x(n_1 - n_2) - \hat{x}(n_1 - n_2))^2 \tag{5 1}$$

$$PSNR(db) = 10 log(255^2)/MSE \tag{5 2}$$

5 0 3 comparison with JPEG

The performance of this coder is compared to a widely available version of JPEG AC JPEG does not allow the user to select a target bit rate but instead allows the user to choose a "Quality factor " While there is some loss of resolution in both there are noticeable blocking artifacts in the JPEG version From the table no 5 and 6 it is shown that the results of embedded coder is better than JPEG AC

Another interesting property of the embedded coding is that even at extremely high compression ratios,the image is recognizable At a compression ratio of 512 1 the image quality of Lenna is poor but still recognizable This is not the case with conventional block coding schemes, where at such high compression ratios there would be insufficient bits to even encode the DC coefficients of each block The un avoidable artifacts produced at low bit rates using this method are typical of wavelet coded schemes coded to same PSNR's However subjectively they are not nearly as objectionable as blocking effects typical of block transform coding schemes

An interesting and perhaps surprising property of embedded coding that has been observed is that when the encoding or decoding terminated during the middle of a pass or in the middle of the scanning of a subband , there are no artifacts produced

that would indicate where the termination occurs

5 0 4 comparison with wavelet based technique

Another interesting figure of merit is the number of significant coefficients retained DeVore et al used wavelet transform coding to progressively encode the same im age[6] Using 68272 bits, (8534 bytes, 0 26 bpp), they retained 2019 coefficients and achieved rms error of 15 30 (MSE = 234, 24 42 db), where as using embedded scheme 9248 coefficients are retained, using 8192 bytes The PSNR of these two examples differ by 8 db Part of the difference can be attributed to the fact that the Harr basis was used in [6]

5 0 5 Limitation of present encoder

The primary drawback of this technique is that because of multiple passes required the algorithm tends to run slowly For the 256 X 256 LENNA at 0 25 bits/pixel and 0 5 bits/pixel the encoder and decoder took 0 45 and 0 65 seconds respectively For the 512 X 512 LENNA at 0 25 bits/pixel the encoder and decoder took 2 seconds on HP 9000/735 machine(32 bit processor)

5 0 6 A comparison to vector quantizer

Vector quantization [5] has been shown to be an effective tool for coding subband coefficients This is because of its ability to exploit redundancies between subbands and because the rate distortion characteristics of vector coding is always better than the equivalent scalar implementation However vector quantization requires the de velopment of code books and the computational burden of choosing the appropriate vectors to represent the data

The results show that the method implemented here gives a variable means of compressing the coefficients of a wavelet transformation in a tree structured decompo sition, without the use of vector quantization This new technique offers a procedure involving relatively straight forward segmentation and coding which can be accom plished with much less computational burden than the full search vector quantization

technique Though, comparable vector quantization techniques have better shown to produce high compression ratios with very good quality, but as this method of successive approximation coupled with adaptive arithmetic coding shows, a scalar method taking advantage of correlation between subbands can also produce good compression with much less computational effort In addition, we are saved the training costs associated with a setting up of code book to begin with

TABLE 5
Coding results for JPEG with arithmetic coding
LENNA 256 by 256

bit rate	PSNR
0 124	23 27
0 291	27 49
0 516	30 21
0 762	32 18

TABLE 6

Embedded
Coding results for [redacted] coder
LENNA 256 by 256

bit rate	PSNR
0 098	21 28
0 27	28 13
0 5	31 2
0 85	34 15
1 2	36 08
2 5	42 11

TABLE 7

Embedded
Coding results for [redacted] coder
BABOON 256 by 256

bit rate	PSNR
0 0675	18 9
0 125	25 20
0 25	28 13
0 5	32 56
1	34 7

TABLE 8

Embedded
Coding results for [redacted] coder
NUTAN 256 by 256

bit rate	PSNR
0 0675	15 26
0 125	16 2
0 25	23 025
0 5	29 04
1	33 7

TABLE 9

Coding results for 512 by 512 LENNA showing peak signal to noise (PSNR) and the number of wavelet coefficients that were coded as nonzero

bytes	rate	compression	PSNR(db)	successive coeff
512	0 015625	512 1	22 124	603
1024	0 03125	256 1	24 06	1210
2048	0 0625	128 1	25 225	2358
4096	0 125	64 1	29 46	4641
8192	0 25	32 1	32 45	9248
16384	0 5	16 1	36 76	18753
32768	1	8 1	40 75	39320

TABLE 10

Codıng results for 512 by 512 BABOON showıng peak sıgnal to noıse (PSNR) and the number of wavelet coefficıents that were coded as nonzero

bytes	rate	compressıon	PSNR(db)	sıgnıf coeff
512	0 015625	512 1	18 22	552
1024	0 03125	256 1	21 798	1126
2048	0 0625	128 1	22 13	2240
4096	0 125	64 1	24 72	4943
8192	0 25	32 1	27 72	10640
16384	0 5	16 1	30 76	21773
32768	1	8 1	35 75	46499

Chapter 6

CONCLUSION AND FUTURE SCOPE OF WORK

A new technique for image coding has been implemented that produces fully embedded bit stream. Furthermore, the compression performance of this algorithm is competitive with virtually all known techniques. The remarkable performance is attributed to the use of following features:

a discrete wavelet transform, which decorrelates most sources fairly well

zerotree coding, which by predicting insignificance across scales provides substantial coding gains

successive approximation, which allows the coding of multiple significance maps using zerotrees and allows the encoding or decoding to stop at any point

adaptive arithmetic coding, which allows the entropy coder to incorporate learning into the bit stream itself

The precise rate control that is achieved with this algorithm is a distinct advantage. The user can choose a bit rate and encode the image to exactly the the desired bit rate. Furthermore since no training of any kind is required the algorithm is fairly general and performs remarkably well with most off the images.

6 1 Future scope of work

This EZW algorithm can be extended to video where in addition to spatial redundancy, temporal redundancy can also be eliminated, using standard motion estimation and compensation algorithms to get further compression

References

1 S Mallat " A Theory for Multiresolution signal decomposition The Wavelet representation " IEEE Trans Pattern Anal Mach Intell vol 11 july 1989

2 S Mallat " Multi channel decomposition of images and wavelet models ' IEEE Trans Accoust Speech and Signal processing Vol 37 dec1990

3 Ianh Witten, Radford M Neal, and John G Cleary, Arithmetic coding for data compression " Comm ACM, VOL 10 june 87

4 Mark R Bauham Barry J Sullivan A Wavelet transform image coding technique with a quadtree structure Int conf on ASSP 1992 V 653

5 M Antonini, M Barland, P Mathiew, Daubechies ' Image coding using wavelet transform " IEEE trans on image processing vol1 apr1992

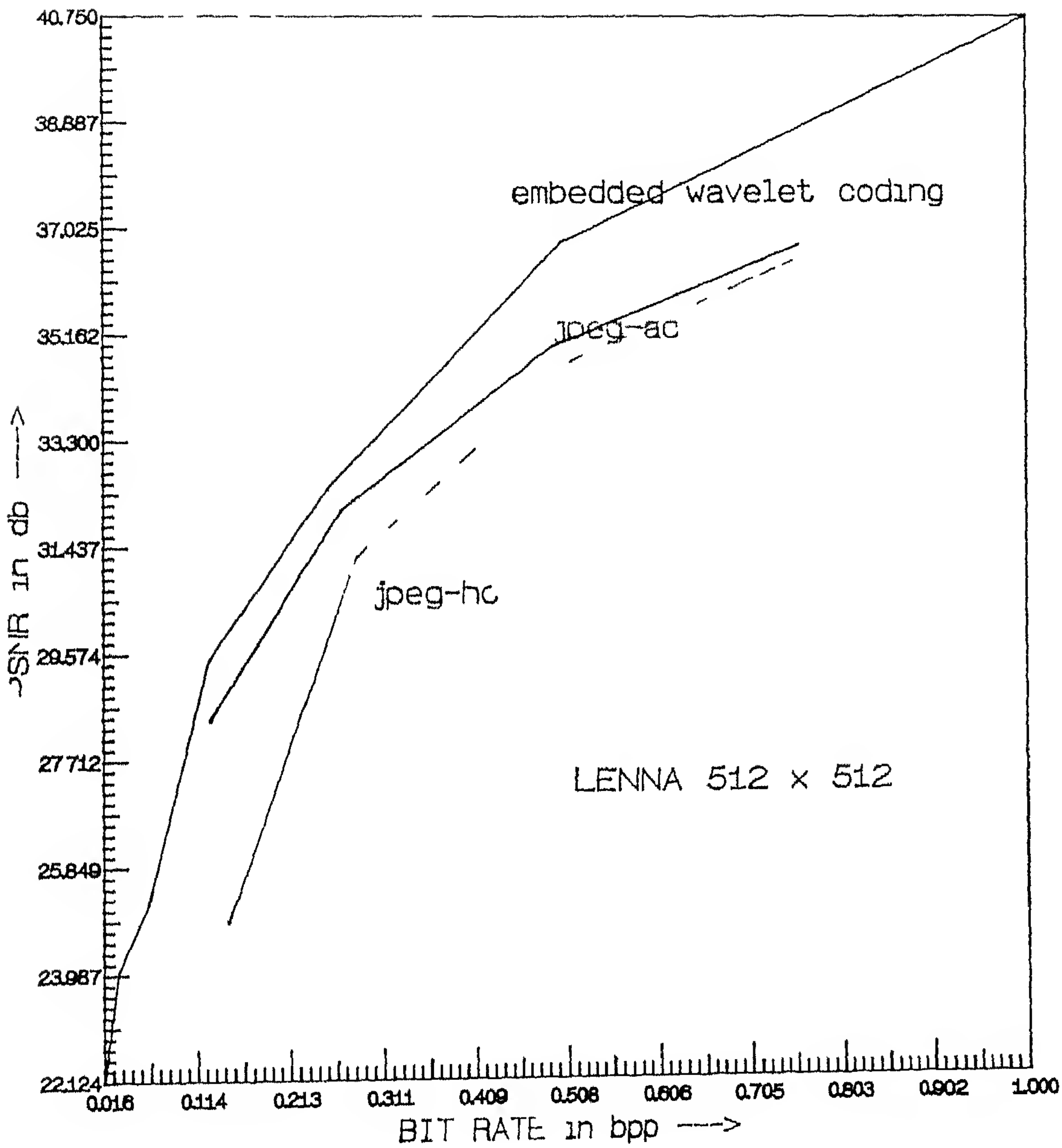
6 R A DeVore B Jaweith and B J Lucier ' Image compression through wavelet transform coding" IEEE Trans inform theory, vol 38, march 1992

7 J M Shapiro " An embedded Hierarchical Image coder using zerotrees of wavelet coefficients" IEEE Trans signal processing vol 41 dec 1993

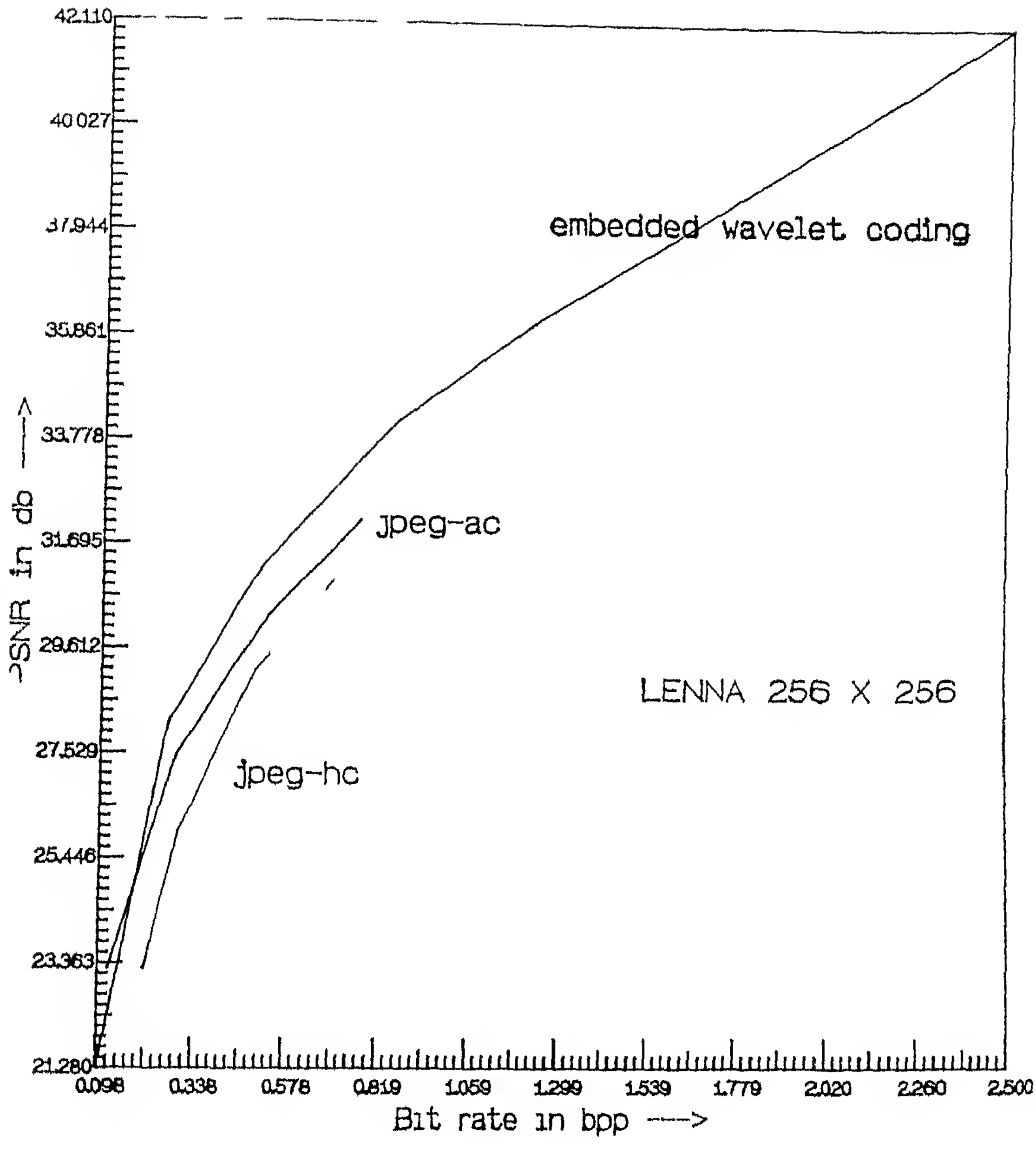
8 I Daubechies Orthonormal bases of compactly supported wavelets' IEEE Trans inform theory, vol 36, sep 1990

9 Rabbani ' Digital image processing'

10 Y T Chan Wavelet basics '



40.750
38.887
37.025
35.162
33.300
31.437
29.574
27.712
25.849
23.987
22.124
0.016
0.114
0.213
0.311
0.409
0.508
0.606
0.705
0.803
0.902
1.000
PSNR in db ---->
BIT RATE in bpp ---->
embedded wavelet coding
jpeg-ac
jpeg-hc
LENNA 512 x 512



42.110
40.027
37.944
35.861
33.778
31.695
29.612
27.529
25.446
23.363
21.280
PSNR in db --->
embedded wavelet coding
jpeg-ac
jpeg-hc
LENNA 256 X 256
0.098
0.338
0.578
0.819
1.059
1.299
1.539
1.779
2.020
2.260
2.500
Bit rate in bpp --->

1 16

1 32

1 64

1 5 12

Image of NUTAN 256 X 256 at different compression ratios

Image of Baboon 256 X 256 at different compression ratios